

Reconstructing Population-scale Exposures from Dose Biomarkers using Bayesian Inference

Michael D. Sohn¹, Thomas E. McKone^{1,2}, Jerry N. Blancato³

¹Lawrence Berkeley National Laboratory
Environmental Energy Technologies Division
Indoor Environment Department
Berkeley, California 94720

²School of Public Health
University of California
Berkeley, California 94720

³National Exposure Research Laboratory
United States Environmental Protection Agency
Las Vegas, Nevada 89193

This work was supported in part by the U.S. Environmental Protection Agency, National Exposure Research Laboratory through Interagency Agreement #DW-988-38190-01-0 and carried out at Lawrence Berkeley National Laboratory through the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Reconstructing Population-scale Exposures from Dose Biomarkers using Bayesian Inference

Michael D. Sohn¹, Thomas E. McKone^{1,2}, Jerry N. Blancato³

¹Lawrence Berkeley National Laboratory, Berkeley, California 94720

²University of California, Berkeley, California 94720

³National Exposure Research Laboratory, United State Environmental Protection Agency, Las Vegas, Nevada 89193

submitted to Journal of Exposure Analysis and Environmental Epidemiology

LBNL Report: 50588

January 7, 2003

This document is a preliminary draft. It has not been formally released by the United States Environmental Protection Agency and should not at this stage be construed to represent Agency policy. It is being circulated for comments on its technical merit and policy implications.

Abstract

Physiologically based pharmacokinetic (PBPK) modeling is a well-established toxicological tool designed to transform an exposure into a target tissue dose. The emergence of federal and state programs for environmental health tracking and the availability of exposure monitoring through biomarkers creates the opportunity to apply PBPK models to reconstruct population exposures to environmental contaminants from urine, blood, and tissue samples. In this paper we lay out and illustrate a plan for examining population exposures using an integrated Bayesian statistical framework. The approach provides flexibility for evaluating multiple exposure scenarios and alternative datasets, which will be critical for systematically reconstructing exposures using biomarkers from a large population. We demonstrate the approach by reconstructing population-scale source-to-dose relationships for a population exposed to trichloroethylene (TCE) through inhalation. We used biomarker data from eight adult males exposed to TCE vapors in air for 240 minutes in an enclosed chamber. In the application, two groups of individuals had distinctly different TCE concentrations in blood despite being contained in the same experimental chamber. We successfully reconstructed the exposure scenarios for both subgroups - although the reconstruction of one subgroup is different than what is believed to be the true experimental conditions. We were however unable to predict with high certainty the concentration of TCE in air. We also present several methods for improving the reliability of the population-scale exposure reconstructions.

Introduction

Physiologically based pharmacokinetic (PBPK) modeling is a well-established toxicological tool designed to transform an exposure into a target tissue dose (Ramsey and Andersen, 1984; Klaassen, 1996). In reviewing the published literature in journals such as *Toxicology and Applied Pharmacology*, *Risk Analysis*, and *Environmental Health Perspectives*, we observe that PBPK modeling is poised to move beyond the phase of model development to the phase of model application. The motivation for such a move is driven in large part by the emergence of federal and state programs for environmental health tracking. There is certainly consensus in the community on how to build PBPK models—basing them on known physiological processes that are easily generalized (blood flow rates, tissues volumes, breathing rates, etc.), on chemical-specific processes that are predicted from existing data and regression models (partition coefficients, chemical density, molecular weight, etc.), and on processes, such as metabolic constants that are highly variable among species and individuals (Gargas et al., 1989). Moreover, the US EPA National Exposure Research Laboratory in Las Vegas has had a generalized PBPK model called the Exposure Related Dose Estimating Model (ERDEM) for some five years.

But there is less consensus or discussion on how to use PBPK models to find environmental determinants of chronic disease from biomonitoring - such as measured pollutant levels in blood and urine samples for a cross section of the population. Yet such applications of PBPK models are needed to build hypotheses about possible relationships between exposures, dose, and disease, to monitor trends in environmental quality and disease, and to provide public health professionals with reliable information

for early detection and prevention of diseases. Rather than new models, such issues demand a new framework for applying PBPK models.

An opportunity for such a new model framework is improving the process of relating the production, use, or release of a chemical and the corresponding dose to an exposed population. Most earlier PBPK applications were based on or applied to well-understood or characterized studies of an individual or a small cohort. The new challenge is how to apply PBPK models to larger and more poorly characterized human populations which have highly variable exposures, activities, physiology, and pharmacokinetics. An important research question here is whether PBPK models are broadly applicable as tools for relating dose biomarkers to measures of population exposure and health risk. If feasible this method offers the opportunity to better relate biomarkers to specific sources of exposure, e.g., household pesticide use versus food residues, and VOC emissions from consumer products versus those from automobiles or from stationary sources.

Although limited to date, there have been recent advancements on this front. The EPA dioxin reassessment used PBPK models to evaluate the reasonableness of their earlier estimated cumulative dietary intake of dioxin compounds (Pinsky and Lorber, 1998; USEPA, 2001). Wallace and Pellezzari (1995) and Wallace (1997) assessed the utility of using exhaled breath for estimating exposure and body burden for volatile organic compounds based on PBPK models. Chinnery and Gleason (1993) and McKone (1993) used PBPK models of chloroform applied to breath samples reported by Jo et al. (1990) to determine the relative contribution of inhalation and dermal exposure routes for adults showering with water containing residual chloroform for disinfection. And by developing methods to treat PBPK model parameters as random variables within the

constraints of empirically observed distributions, Bois et al. (1996a, 1996b) and Gelman et al. (1996) built population-based models for tetrachloroethylene and benzene.

Resources and opportunities to produce population-scale source to dose exposure reconstructions may come from new and ongoing national and regional epidemiological surveys. The Center for Disease Control (CDC) National Health and Nutrition Examination Survey (NHANES) (CDC, 2001) employs a home interview with health tests to collect information about the health and diet of people in the United States. It includes data on blood levels of cadmium, lead, mercury, pesticides, and combustion products. Through the National Human Exposure Assessment Survey (NHEXAS) (Sexton et al., 1995), the Children's Total Exposure to Persistent Pesticides and other Persistent Organic Pollutants (CTEPP) (USEPA, 2002), and other programs, the U.S. EPA is developing databases on exposures of human populations to a wide range of pollutants in air, water, food, soil, and indoor/residential environments, and over a wide range of space and time scales. The University of California, Berkeley Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) (Castorina et al., 2002) is collecting biomarkers for pesticides and other important pollutants from mothers and their newborn children in farming communities. The information is gathered from conception through early childhood.

However, biomarkers obtained from these surveys are inherently variable owing to the inter- and intra-individual variability among exposures to the population and the physiology of the individuals in the population. The key questions are whether and how well we can quantify the source to dose relation against the noise contributed by these other factors. Examining the input and output information obtained in these surveys will be critical. What exposure information is currently available, and what

additional information is likely or practical to be obtained? How do we make PBPK models compatible with the available information instead of using the models in the form originally developed – based on controlled laboratory conditions?

Answering these questions is an important direction for ongoing development and application of PBPK models. It will require more use and development of statistical and other quantitative methods for integrating uncertainties and variability in both model predictions and biomarker data.

In this paper we lay out and illustrate a plan for examining these questions using an integrated Bayesian statistical framework. The approach provides flexibility for evaluating multiple exposure scenarios and alternative datasets, which will be critical for systematically reconstructing exposures using biomarkers from a large population. The Bayesian framework will also help us apply PBPK models for setting priorities for exposure and health monitoring programs, and for directing what exposure information to gather in the near term.

We demonstrate the approach by reconstructing population-scale source-to-dose relationships for a population exposed to trichloroethylene through inhalation. We also discuss methods to determine what exposure assessment information is important in the reconstructions, what information other than that commonly gathered in exposure assessments could improve the dose reconstruction, and how quality and quantity of data affects the reconstructions, thereby assisting future exposure and epidemiological studies.

Method: Exposure Classification Using Bayesian Statistics

In the most general sense, an exposure assessment involves quantifying a link between a source of contamination, its transport and transformation among a set of environmental media, human contact with exposure media, and the route of application or entry (USEPA, 1989; McKone and Daniels, 1991a; USEPA, 1992; Zartarian et al., 1997). Environmental media include outdoor air, indoor air, ground-surface soil, root-zone soil, plants, ground water, and surface water in a contaminated landscape as well as carpets, furnishings, etc. in indoor environments. Exposure media include substances with which we have direct contact such as outdoor air, indoor air, food, household dust, household surfaces, homegrown foods, animal food products, and tap water. Exposure pathways define the links between an environmental media and exposure media for inhalation, ingestion, and dermal uptake routes of exposure. Potential dose, expressed as average daily dose, is the amount of material per unit of body weight per day (mg/kg-d) that enters the lungs (inhalation route), enters the gastrointestinal tract (ingestion route), or crosses into the stratum corneum (dermal-contact route) (USEPA, 1989; McKone and Daniels, 1991a). This total potential dose is commonly used as a basis for projecting the incidence of health detriment within the exposed population.

However, each component in the exposure-to-dose link includes some level of uncertainty or bias. For example conceptual and process models in PBPK models are developed from infrequently sampled yet highly variable and uncertain data. Ignoring the variability and uncertainty in the models can imply over-confidence in the PBPK model, and can cause erroneous estimates of exposure-to-dose relationships.

We confront these uncertainties using Bayesian inference methods (see e.g., Morgan and Henrion, 1990). In this approach, the practitioner first develops

mechanistic, statistical, and/or empirical models that predict the source to dose relationship. Any unknown, uncertain, or variable model input is probabilistically described using parametric or non-parametric uncertainty distributions. Examples of unknowns include the alternative exposure scenarios, variability in the pharmacokinetics, alternative conceptual models such as two-compartment or five-compartment PBPK models, first-order or second-order environmental degradation, well-mixed or multi-compartment indoor air models, and various model parameter uncertainties. Field data, epidemiological studies, best engineering judgement, and any quantitative or subjective information are possible sources for developing value ranges of the probabilistic distributions. Generally, the practitioner will assign wide uncertainty distributions due to the limited information.

The practitioner next predicts model endpoints that can be compared to specific biomarker data. A Monte Carlo or Latin Hypercube sampling technique may be applied to generate a library consisting of several thousand realizations of exposure scenarios and biomarker predictions. Sufficient sampling of the uncertainty distributions is essential to represent the full range of possible exposure scenarios. One method for testing sufficiency of sampling is by increasing the sample size until changes in summary statistics (e.g., means, variances, coefficients of variation) of model predictions are negligible with each increase in sample size. The values and the uncertainties of model predictions, model input parameters, and conceptual models employed are often referred to as the "prior". They collectively define the model formulation prior to being compared to data.

The practitioner assesses the agreement between model predictions and biomarker data using a technique called Bayesian updating. Brand et al., 1994, Sohn et

al., 2000, and Bois et al., 1996a describe the technique in detail. Recent applications in multi-pathway, multi-parameter, environmental systems include assessing environmental health risk (e.g. Taylor et al., 1993; Spear and Bois, 1994, Brand et al., 1995; Pinsky and Lorber, 1998), analyzing groundwater monitoring data (e.g. Dilks et al., 1992; Wolfson et al., 1996; Sohn et al., 2000), and conducting environmental value-of-information analyses (e.g. Finkel and Evans, 1987; Dakins et al., 1996). We therefore briefly describe the relevant details of the technique as they pertain to population-scale source-to-dose analyses.

Each dose prediction in the library of model simulations is compared to biomarker data using Bayes' rule (Equation 1).

$$p(Y_k | O) = \frac{L(O | Y_k) p(Y_k)}{\sum_{i=1}^K L(O | Y_i) p(Y_i)} \quad (1)$$

where $p(Y_k | O)$ is the probability of the k^{th} Monte Carlo simulation making prediction Y_k given the biomarker data O ; $L(O | Y_k)$ is the likelihood of observing measurements O given model prediction Y_k ; $p(Y_k)$ is the prior probability of the k^{th} Monte Carlo simulation; and K is the number of Monte Carlo simulations. Before data comparison, each of the model realizations are usually assumed equally likely (i.e., $p(Y_k)=1/K$).

The probability $p(Y_k | O)$ is often referred to as the posterior probability of the k^{th} realization since it describes the probability after the k^{th} realization is compared to data. In this case, the posterior probability describes the degree that the k^{th} dose prediction - and the associated model input parameters, conceptual models, and exposure scenario used to generate that prediction - accurately describes the biomarkers. The posterior probability thus replaces all of the uncertain prior probabilities (e.g., unknown model input parameters and exposure scenarios) and model predictions (e.g., dose estimates).

Brand et al. (1994), Sohn et al. (2000), and Sohn et al. (2002) provide detailed description for estimating posterior means, variances, and correlation coefficients.

The likelihood function, $L(O | Y_k)$, in Equation 1 quantifies the error structure of the data, i.e., the differences between the data and the model predictions resulting from measurement error, spatial and temporal averaging or correlations, and imperfect model representation. If many independent measurements are considered, for example following random samples in a large epidemiological survey, the likelihood of observing all of the measurements is the product of all of the individual likelihoods:

$$L(O | Y_k) = \prod_{s=1}^S L(O_s | Y_{s,k}) \quad (2)$$

where S is the number of independent measurements.

For unbiased measurements with a normally distributed error, a Gaussian likelihood function is appropriate (see e.g., Taylor et al., 1993; Brand and Small, 1995; Dakins et al., 1996; and Sohn et al., 2000). Sohn et al. (2000) suggest alternative likelihood functions when errors are not independent and normally distributed.

We note that the Bayesian updating procedure presented here is executed in several sequential steps: (1) develop models and predict dose, (2) compare predictions to data, and (3) update uncertainties using the likelihoods. Alternatively, one could predict model endpoints and compare them to biomarkers using one of several other variations of Bayesian updating such as Markov Chain Monte Carlo or Gibbs sampling (see e.g., Gelman et al., 1995, 1996; Roy and Georgopoulos, 1998). These approaches apply Equations 1 and 2 iteratively and are capable of quickly searching the model parameter uncertainties. However, we did not find these more complex methods necessary for the illustrative demonstration that follows. Executing the PBPK models

was very fast on a standard desktop personal computer so we did not have difficulty adequately sampling the uncertainties and variability in the model inputs.

Application to a Trichloroethylene Exposed Cohort Study

We demonstrate the approach using data from a group of individuals exposed to TCE in a controlled laboratory setting. We recognize that the variability and uncertainties in a large health tracking study will be larger than those found in this dataset. Nevertheless, the errors and variability in the data from this small sample are sufficient to demonstrate the approach and highlight the difficulties that arise in reconstructing exposures to populations.

The biomarker data come from experiments conducted by Fisher and colleagues (Fisher, 1998). Eight adult males were exposed to (TCE) vapors in air for 240 minutes in an enclosed chamber. The details of the experiment and the laboratory setup are described in Fisher, 1998. Figure 1 plots the times-series concentrations of TCE in the venous blood of the eight exposed subjects. The pooled variance of the log-transformed concentrations (Weisberg, pg. 90) is 0.38 and the error structure of the log-transformed concentrations appears to be Gaussian. All eight individuals were contained in the same chamber, though we note that three individuals, which we call subgroup A, have significantly lower concentrations than the other five individuals, which we call subgroup B. There are no obvious experimental or equipment variations that explain the differences in the blood concentrations.

We first developed a PBPK model to predict the concentration of TCE in venous blood from exposure by inhalation. Ramsey and Anderson (1984) provide details of a typical first-order multi-compartment PBPK model. Researchers at the US EPA successfully developed a multi-compartment PBPK model using the ERDEM model

(formerly referred to as DEEM) for an individual exposed to TCE by inhalation (Blancato et al., 2000). Because this model required information and detail not usually available in population-scale exposure studies, we used a simpler PBPK model for the Bayesian inference process, though we could have developed a simple PBPK model in ERDEM. Our model is a first-order five-compartment model consisting of four well-mixed tissue groups; fat, liver, slowly- and rapidly perfuse tissue, and a pulmonary compartment to represent the blood/air transfer in the lungs. Metabolism in the liver was described by Michaelis-Menton kinetics. The model was coded in FORTRAN and solved using the VODE solver (Brown et al., 1988).

Table 1 summarizes the parameters needed to characterize the uncertainty and variability in the exposure to TCE and the pharmacokinetics of this population. They include wide value ranges for both exposure scenarios and human pharmacokinetics. We predicted TCE concentration in blood by sampling the probability distribution in Table 1 using Latin Hypercube sampling and applying the scenarios to the five-compartment PBPK model. Twenty thousand samples and model simulations sufficiently sampled the parameter value ranges summarized in Table 1. Figure 2 shows the range of predicted TCE concentration in blood. The wide uncertainty bounds in Figure 2 result from the combined uncertainties in the exposure scenario and pharmacokinetics.

We next applied the Bayesian exposure classification algorithm to estimate the exposures for the two subgroups. Figure 3 shows the reconstructed exposure scenarios and Figure 4 shows the reconstructed TCE concentration in the blood. The exposure reconstruction for subgroup A is consistent with the experimental conditions. The uncertainty bounds narrow to the values reported by Fisher (1998) for the duration of

the exposure (Figure 3), the onset of the exposure (Figure 3), and the predicted concentration in blood (Figure 4). The predicted TCE concentration in air (Figure 3), though its median is close to the reported air concentration, still contains considerable uncertainty. This is despite the high temporal resolution of the biomarkers.

In contrast to subgroup A, the exposure reconstruction for subgroup B are not consistent with the values reported by Fisher (1998). The Bayesian updating reduces uncertainty, but the predicted exposure duration and onset are longer than what is reported to have occurred in the experiment.

The differences in reconstructions for the two datasets are in large part due to the unexplained differences in the TCE concentrations. Our discussions with the authors of the experimental study did not provide an obvious explanation. One explanation, though highly speculative, may be that the experimental chamber was not uniformly well-mixed. A non-well-mixed chamber could produce turbulent eddies of high concentration which some of the individuals (i.e., those defined in subgroup B) were exposed to.

Another explanation may be that the two subgroups reflect different metabolism profiles or breathing rates by the subjects. However, these do not explain the presence of two unique groups and not a simple continuum. If metabolism or breathing rates are random effects, and thus probabilistically distributed over some quantifiable range, we would expect the blood concentrations to be somewhat evenly spread throughout the range of concentrations. Instead we find two unique groups. In addition, we can compare the metabolism and breathing rates of the subjects by comparing the slopes of the curves in Figure 1 after exposure ends. We can visually see that the slopes are

similar for each subject, suggesting that the various TCE elimination routes are similar across the population.

Finally, we used Bayesian inference to reconstruct the exposure of the whole population. Figure 3 and 4 shows the reconstructed exposure conditions and the predicted TCE concentration in blood. Based on the reported conditions of the experiment, the predicted TCE concentration in blood (Figure 4c) has erroneous uncertainty bars. They are so narrow that a large percentage of the data fall outside of the two-sided 95% confidence interval. This problem arises because the exposure reconstruction assumed a single-mode population, but the data clearly show two distinct modes (or subgroups). We show this result to emphasize the importance of applying correct modeling assumptions when developing source-to-dose links, irrespective of the exposure reconstruction method employed. In this case, a first-order PBPK model coupled to a single-zone indoor air chamber is inappropriate for a dataset that is bi-modal.

In actual health-tracking we could not expect to have sufficient data to disaggregate the population into subgroups, or to even realize that subgroups exist. We may have more sampled individuals, but it is unlikely that we would have such high temporal resolution for many individuals. In Figure 5, we constructed an example of the biomarker data that might come from a true health-tracking study. We constructed this figure by randomly selecting data at various times from the study population. Here we see that random selection of observations from this group tends to mask the two-modes of the population. The time series in Figure 5 appears to represent a single-mode population with error due to random inter-individual variability.

Figure 6 shows the exposure reconstruction for this smaller dataset. The uncertainty around the predicted TCE concentrations, though wide, appears to correctly bound the data. The predicted exposure conditions are also wide owing to the error in the data, the possible non-well mixed conditions in the chamber, and uncertainty in the pharmacokinetics (Table 1).

Discussion

In the preceding paragraphs, we applied Bayesian inference tool to reconstruct exposure conditions. The process is not straightforward and can be confounded by heterogeneity and variability in exposure conditions and metabolism. Though the results suggest that two sub-groups existed, and that they were exposed to different air concentrations, there still remained considerable uncertainty in the exposures estimates. For example, we were unable to predict with high certainty the concentration of TCE in air from any of the datasets. An important limitation of using PBPK models to interpret biomarker data may be the non-uniqueness of inverse solutions due to the combined effects of variability in exposures and human pharmacokinetics.

Nevertheless, the Bayesian exposure assessment approach offers some key advantages for reconstructing exposures. An important attribute of the approach is its flexibility for analyzing and comparing the utility of various types or quantities of data without excessive computational or numerical burdens. For example, we recalculated exposure reconstructions using data from subgroup A only, B only, A and B combined, and a randomly sampled set without re-executing Monte Carlo simulations of the PBPK model.

We can also test whether more data or other types of data could improve the exposure reconstructions. For example, could better information about the environmental, exposure, and pharmacokinetic properties improve reliability of exposure reconstruction, and at what cost? As a demonstration of such an exploratory search, we recalculated each of the exposure reconstructions with the added assumption that (1) the exposure duration is known or (2) the time of exposure onset is known. Including the additional information did not reduce the uncertainty in the predicted concentration of TCE in air in any of the reconstructions; we therefore did not plot the results. This suggests that the population-scale variability of the pharmacokinetics alone or in correlation with other unknowns dominate the uncertainty of the predicted TCE concentration in air. One should thus not expend excessive resources obtaining either information alone if reducing uncertainty in the predicted concentration of TCE in air is the primary objective.

Recommendations

Given the potential problems we observe for exposure reconstruction, we must consider what properties of a biomarker or other types of information can improve the reliability of the exposure reconstruction process. Because the dose delivered to an exposed individual depends on (a) the time scales of the pharmacokinetics of the agent, (b) the route of entry, and (c) the rate of intake or uptake at the human/environment boundaries, we must first recognize the importance of selecting the most appropriate time scale for collecting information. For example, resolving the temporal variability of exposure events requires differentiating between (i) a recent relatively mild peak exposure or a long-term relatively high exposure, or (ii) many common exposures occurring simultaneously. How persistent must the biomarker (body burden) be

relative to exposure duration for this task? That is, if we want to infer exposure to a pollutant over a one-week period, we must consider the minimum biological half-life of the biomarker required to reconstruct doses reliably.

Of similar importance is the time characteristics of PBPK model input parameters. For example, breathing automatically averages air concentrations over the duration of a breath (at least at the lung level), and drinking and eating are discontinuous. These conditions define both limiting processes and the time-averaging periods for health-relevant doses. PBPK models establish the characteristic time of pollutant within the human body. This time is needed to classify exposures as intermittent (going to zero or negligible levels periodically or randomly) versus continuous (with various degrees of stability/uncertainty), and as sequential, additive, or cumulative. Exposures to carcinogens at low rates of uptake (when the cumulative damage rate is proportional to the uptake) require a dose assessment with a time resolution that need only reflects the cumulative uptake or intake of the agent into the body. In contrast, an agent such as an acid gas (where short term non-linear effects with large variations in respiratory susceptibility are important) requires much more detailed specification of the time, population, and even spatial resolution of exposure. PBPK modeling may help in the design, timing, and placement of measurements that are necessary for developing such a technique.

Based on the results presented here, we propose that the persistence of the biomarker should be long relative to exposure duration for estimating long-term, or population scale, exposure effects. That is if one wants to infer exposure to a pesticide over a one week period, it is useful to have a biomarker that persists for more than a week. Perhaps the best marker is one that is truly cumulative, i.e. a chromosome

aberration that is heritable from one cell generation to another. But how do we establish a lower bound on biomarker persistence? In the TCE dataset, high temporal resolution throughout the exposure event allowed us to reconstruct exposures for each of the subgroups with reasonable success. Could even better resolution or a more persistent biomarker improve the predictions of the TCE concentration in air?

Better understanding of how variability in metabolism impacts the reliability of PBPK models to determine unique links between exposure and tissue dose is also critical when designing exposure classification studies. Are PBPK models so limited by metabolic pathway uncertainty/variability, that higher compartment resolution – and thus the precision of the model predictions or exposure estimates – is of no value? This could explain some of the uncertainty in the exposure reconstructions for the various combinations of the TCE dataset. Or will age-specific variation of physiological parameters, particularly for persistent chemicals such as DDT, PCBs, TCDD that can accumulate in tissues (fat in particular) over decades make population-wide exposure-to-dose impact assessments too specific to certain subpopulations, in this case according to age breakdown?

Such exploratory studies and classification studies should be carried out before and during health-tracking studies to ensure that the practitioner obtains the most informative data, whether they are environmental, biological, or chemical (i.e., the properties of the pollutant). However, these data for a wide number of chemicals and exposure routes are not adequately described in the current literature on PBPK models or epidemiological surveys.

As a last point, we note that the reconstruction of the combined dataset (Figure 5) stresses the importance of appropriately developing and applying PBPK models that

are consistent with the data. For example, reconstructing a two-mode population using a single-model PBPK model resulted in incorrect uncertainty estimates. Either the subgroups should be analyzed separately or a two-mode model should have been developed. When the dataset is sparse (Figure 7), the practitioner may satisfactorily describe the data with a single-mode PBPK model.

Concluding Remarks

Health assessments require public-health tracking and health tracking requires exposure tracking, a process for linking body burdens as reflected in biomarkers to a distribution of population doses. Of particular importance is the ability to classify and link health outcomes with exposures to harmful substances—pesticides, industrial chemicals, combustion products, consumer products, etc. This requires both sufficient and reliable information about population exposures and doses to those pollutant sources that most significantly contribute to observed markers of potential health detriment. To make better use of body burden/biomarker data in the process of public health tracking, two essential scientific research tools, models and measurements, must be better integrated. Models provide the means to integrate and interpret measurements, design hypothesis-driven experiments, and predict the effectiveness of risk management strategies. Measurements, in turn, provide tests of the models and “ground truth.”

We presented an integrated approach for improving the communication between the needs and capabilities of modeling with the needs and capabilities of health surveys. The Bayesian statistical approach allows for easily quantifying the value of biomarkers, or other environmental, chemical, biological parameters, for exposure classification. The

analysis of the TCE dataset demonstrated some of the attributes of the approach and, perhaps more importantly, highlighted the difficulties with back-estimating exposures to a large and diverse population. We provide in the discussion and recommendation sections a list of factors that contribute to these difficulties and recommend several quantifiable measures that should be studied before and during health and exposure surveys to identify what types of biomarkers to gather, when to gather them, and how much data is needed.

Acknowledgements

This work was supported in part by the US Environmental Protection Agency National Exposure Research Laboratory through Interagency Agreement #DW-988-38190-01-0 and carried out at Lawrence Berkeley National Laboratory through the US Department of Energy under Contract Grant No. DE-AC03-76SF00098. We thank M. Rigas and F. Power from the US EPA for their help in developing the PBPK model. We thank M. Macleod and W. Riley for their helpful comments on an earlier version of this manuscript.

References

- Blancato, J, F Power, J Fisher and E Kantor (2002) "Application and Use of Dose Estimating Exposure Model (DEEM) for Dose Comparisons After Exposure to Trichloroethylene (TCE)," 39th Annual Meeting, Society of Toxicology, Philadelphia, PA, March 19-23.
- Bogen, KT and Spear, RC (1987) "Integrating Uncertainty and Interindividual Variability in Environmental Risk Assessment," *Risk Analysis* 7(4):427-436.

Box, GEP and Tiao, GC (1973) *Bayesian Inference in Statistical Analysis*, New York, Wiley Classics.

Bois, FY, A Gelman, J Jiang, DR Maszle, L Zeise and G. Alexeef (1996a) "Population Toxicokinetics of Tetrachloroethylene," *Archives of Toxicology* **70**:347-355.

Bois, FY, ET Jackson, K Pekari and MT Smith (1996b) "Population Toxicokinetics of Benzene," *Environmental Health Perspectives* **104**:1405-1411.

Brand KP and MJ Small (1995) "Updating Uncertainty in an Integrated Risk Assessment – Conceptual Framework and Methods," *Risk Analysis*, **15**(6):719-731.

Brown, PN, GD Byrne and AC Hindmarsh (1989), "VODE, A Variable-Coefficient ODE Solver," *SIAM Journal on Scientific and Statistical Computing* **10**(5):1038-1051.

Castorina, R, A Bradman, TE McKone, DB Barr, ME Harnly and B Eskenazi (2002), "Assessing Cumulative Organophosphate Pesticide Exposure and Risk among Pregnant Women Living in an Agricultural Community," *Environmental Health Perspectives* (submitted).

Centers for Disease Control and Prevention (CDC, 2001), "National Report on Human Exposure to Environmental Chemicals," published by and available from the Centers for Disease Control and Prevention, National Center for Environmental Health, Division of Laboratory Sciences, Atlanta, Georgia, www.cdc.gov/nceh/dls/report.

Chinnery, R and K Gleason (1993), "A Compartment Model for the Prediction of Breath Concentration and Absorbed Dose of Chloroform after Exposure while Showering," *Risk Analysis* **13**:51-62.

Jo, WK, CP Weisel and PJ Lioy (1990), "Chloroform Exposure and Body Burden from Showering with Chlorinated Tap Water," *Risk Analysis* **10**:575-580.

Fisher, J (1998), "Physiologically Based Pharmacokinetic Models for Trichloroethylene and its Oxidative Metabolites," an EPA Report draft 2, April 7, 1998 from the Operational Toxicology Branch, Human Effectiveness Directorate, Wright-Patterson AFB, OH 45433.

Gargas, ML, RJ Burgess, DE Voisard, GH Cason and ME Andersen (1989), "Partition Coefficients of Low-molecular Weight Volatile Chemicals in Various Liquids and Tissues." *Toxicology and Applied Pharmacology* **98**:87-99.

Gelman, A, FY Bois and J Jiang (1996), "Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions," *J. American Statistical Association* **91**:1400-1412.

Gelman, A, JB Carlin, HS Stern and DB Rubin (1995) *Bayesian Data Analysis*. Chapman & Hall/CRC; Boca Raton , Florida.

Klaassen, CD (1996), "Distribution, Excretion and Absorption of Toxicants," *Cassarett and Doul's Toxicology, the Basic Science of Poisons*. C. D. Klaassen, M. O. Amdur and J. Doul. New York, NY, Pergammon Press:pp. 33-63.

McKone, TE (1993), "Linking a PBPK Model for Chloroform with Measured Breath Concentrations in Showers: Implications for Dermal Exposure Models," *Journal of Exposure Analysis and Environmental Epidemiology* **3**:339-365.

Morgan, MG and M Henrion (1990), *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press. New York, NY.

Pinsky, PF and MN Lorber (1998), "A model to Evaluate Past Exposure to 2,3,7,8-TCDD," *Journal of Exposure Analysis and Environmental Epidemiology*, **8**(2):187-206.

Ramsey, JC and ME Andersen (1984). "A Physiologically-based Description of the Inhalation Pharmacokinetics of Styrene in Rats and Humans," *Toxicology and Applied Pharmacology* **73**:159-175.

Roy, A and PG Georgopoulos (1998), "Reconstructing week-long exposures to volatile organic compounds using physiologically based pharmacokinetic models," *Journal of Exposure Analysis and Environmental Epidemiology* **8**(3):407-422.

Sexton, K, DE Kleffman and MA Callahan (1995), "An Introduction to the National Human Exposure Assessment Survey (NHEXAS) and Related Phase I Field Studies," *Journal of Exposure Analysis and Environmental Epidemiology* **5**:3229-232.

Sohn, MD, MJ Small and M Pantazidou (2000), "Reducing Uncertainty in Site Characterization Using Bayes Monte Carlo Methods," *Journal of Environmental Engineering-ASCE*, **126**(10):893-902.

Spear, RC and FY Bois (1994), "Parameter Variability and the Interpretation of Physiologically Based Pharmacokinetic Modeling Results," *Environmental Health Perspectives*, **102**(11):61-66.

US Environmental Protection Agency (USEPA) (2001) "Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds," (DRAFT). *Rpt No. EPA/600/P-001*, US Environmental Protection Agency Office of Research and Development, National Center for Environmental Assessments, Washington D.C. Report available at: <http://www.epa.gov/ncea/dioxin.htm>.

US Environmental Protection Agency (USEPA) (2002) "Children's Total Exposure to Persistent Pesticides and other Persistent Organic Pollutants (CTEPP)," US Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC.

Wallace, LA (1997), "Human Exposure and Body Burden for Chloroform and Other Trihalomethanes," *Critical Reviews in Environmental Science and Technology* **27**(2):113-194.

Wallace, LA and ED Pellizzari (1995), "Recent Advances In Measuring Exhaled Breath and Estimating Exposure and Body Burden For Volatile Organic Compounds (VOCs)," *Environmental Health Perspectives* **103**:95-98.

Table 1: Exposure and pharmacokinetic uncertainty or variability. GM and GSD are abbreviations for geometric mean and geometric standard deviation, respectively.

Model Parameter	Range	Distribution
Exposure		
TCE Conc. in Air (ppm)	25 - 300	Uniform
Exposure Duration (hr.)	1 - 6	Uniform
Onset of Exposure (hr.)	-3 - 3	Uniform
Metabolism		
Vmax (units)	GM: 2.84e-5, GSD: 2.38	Lognormal
Km (units)	GM: 8e-4, GSD: 2.8	Lognormal
Pulmonary Flows		
Q air (l/sec)	GM: 0.108, GSD: 2.16	Lognormal
Q fat (l/min)	GM: 0.3, GSD: 2.3	Lognormal
Q slowly perfuse tissue (l/min)	GM: 0.81, GSD: 2.22	Lognormal
Q rapidly perfuse tissue (l/min)	GM: 3.98, GSD: 2.22	Lognormal
Q liver (l/min)	GM: 1.12, GSD: 2.3	Lognormal
Volume		
V fat (l)	GM: 12.87, GSD: 2.1	Lognormal
V slowly perfuse tissue (l)	GM: 42.3, GSD: 2.06	Lognormal
V rapidly perfuse tissue (l)	GM: 12.94, GSD: 2.18	Lognormal
V liver (l)	GM: 2.18, GSD: 2.08	Lognormal
Partition Coefficient		
P blood/air	GM: 18, GSD: 2.18	Lognormal
P fat	GM: 50.9, GSD: 2.3	Lognormal
P slowly perfuse tissue	GM: 1.5, GSD: 2.28	Lognormal
P rapidly perfuse tissue	GM: 3.67, GSD: 2.22	Lognormal
P liver	GM: 5.81, GSD: 2.3	Lognormal

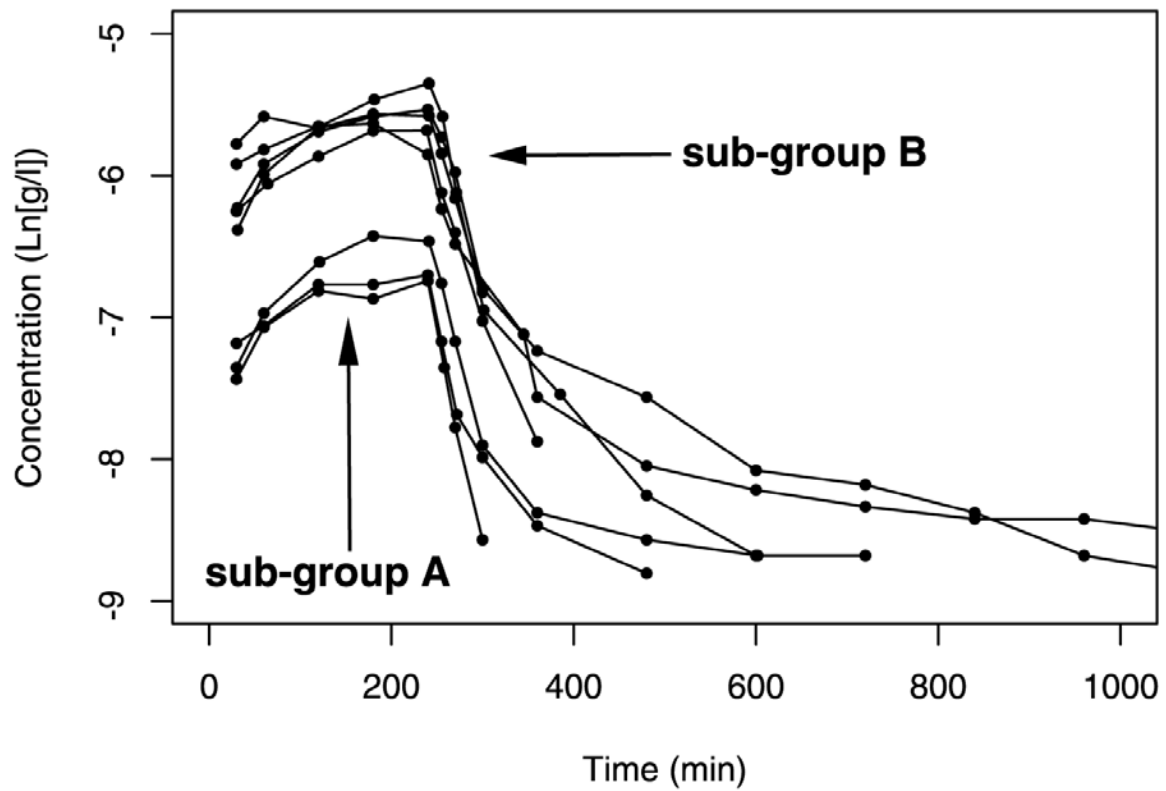


Figure 1: TCE concentrations in venous blood. Eight adult males were exposed in a chamber to TCE in air at a concentration of 100 ppm for 240 minutes. The TCE concentration in air was zero thereafter ($t > 240$ minutes).

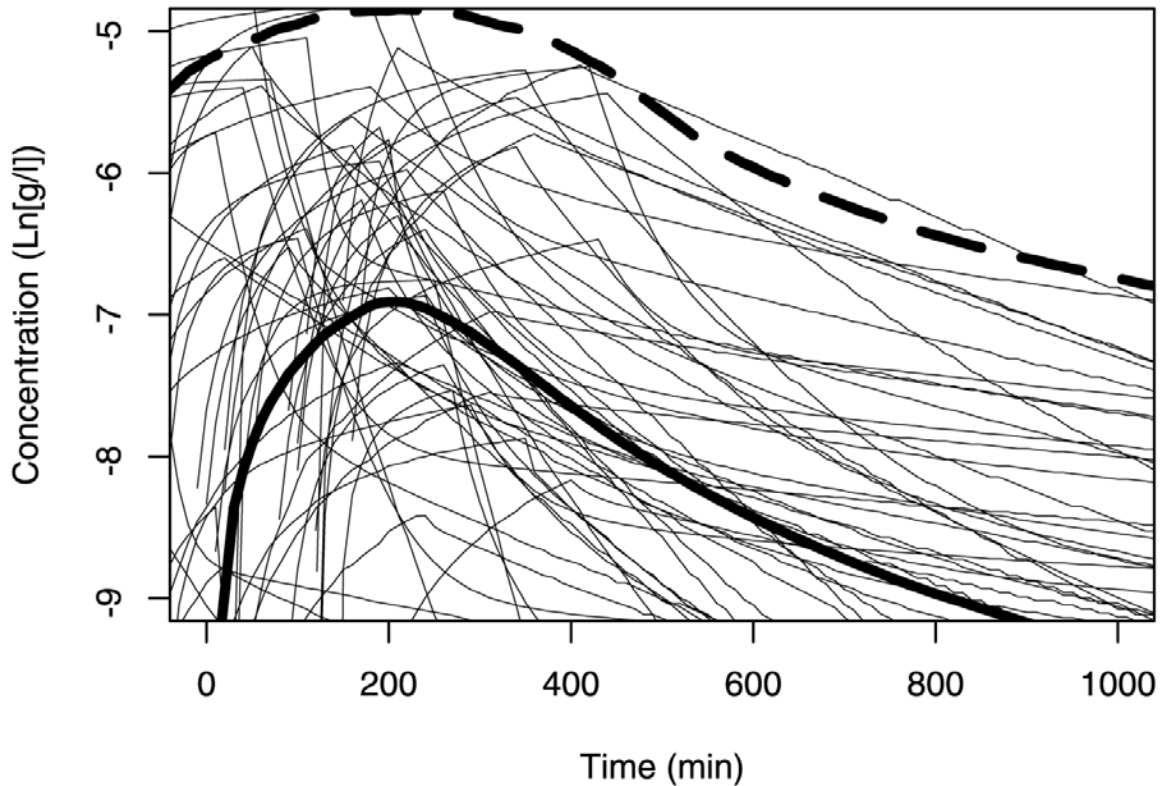


Figure 2: Predictions of TCE concentration in venous blood before comparison to biomarker data. Each thin line represents a concentration profile predicted from a sample of the exposure and pharmacokinetic unknowns in Table 1. Fifty of the 20,000 simulations are plotted here. The thick line is the median of the 20,000 simulations and the thick dotted line is the upper bound of the two-sided 95% confidence interval. The lower bound is below the limits of the y-axis so is not visible on this figure.

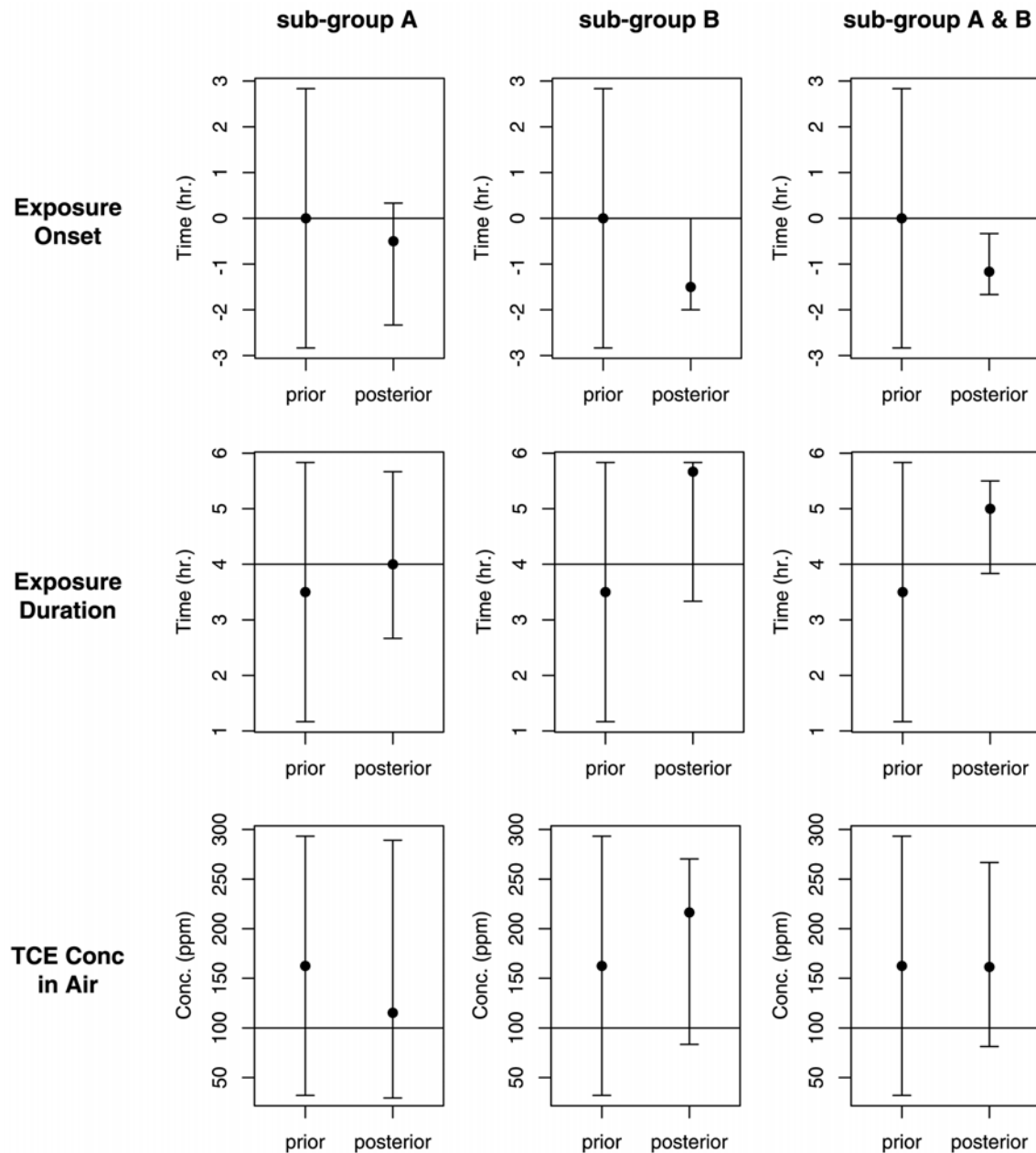


Figure 3: Reconstructed exposure and concentration of TCE in air for the datasets identified in Figure 1. The prior and posterior bars presents the uncertainty before and after comparison to biomarker data, respectively. The whiskers of the bars are the two-sided 95% confidence interval and the circle is the median. The horizontal line represents the value reported by Fisher (1998).

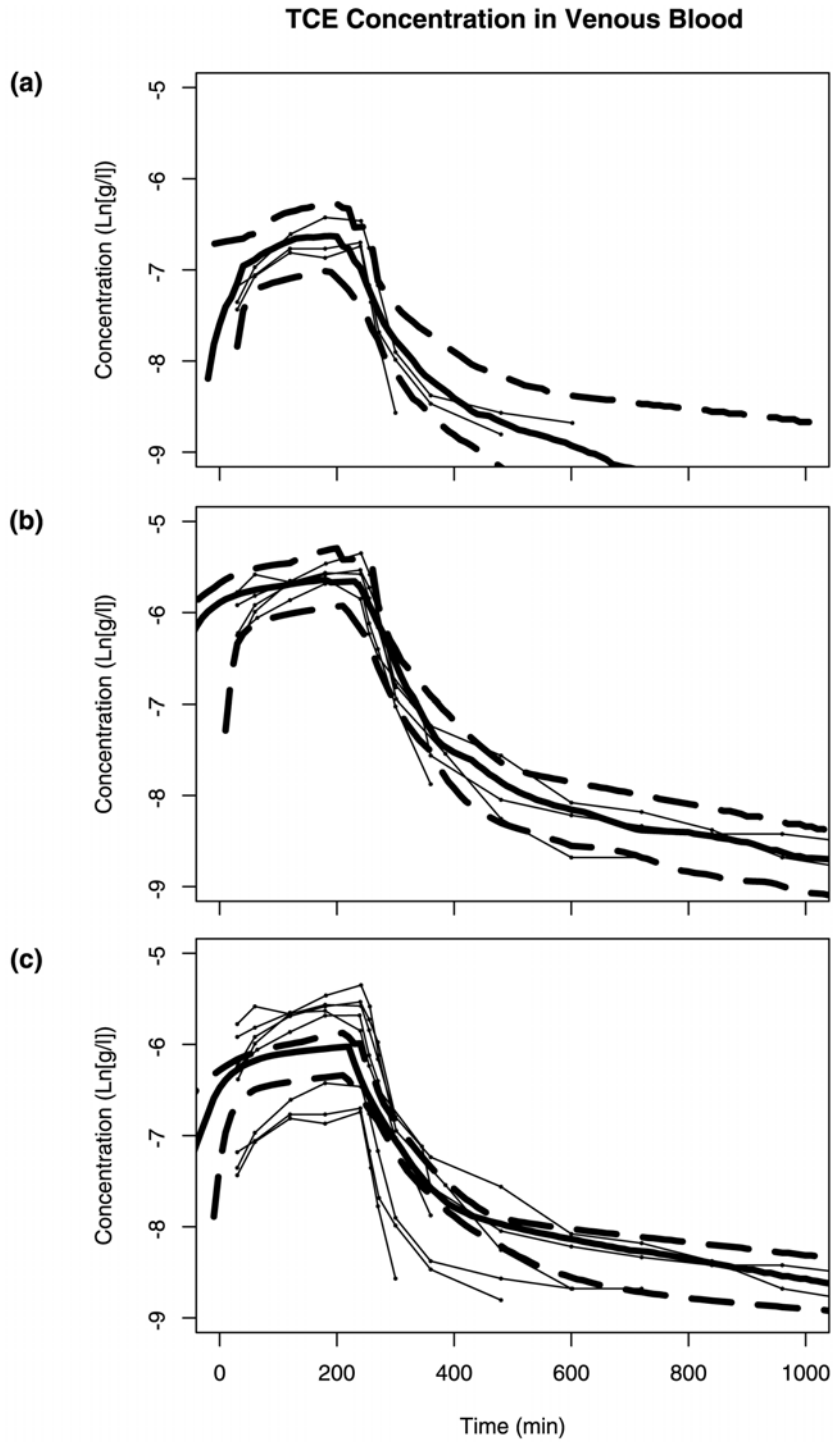


Figure 4: Updated predictions of TCE in venous blood using biomarker data from (a) sub-group A, (b) sub-group B, and (c) sub-groups A and B (see Figure 1). The thick dotted lines define the two-sided 95% confidence interval. The thick solid line is the median.

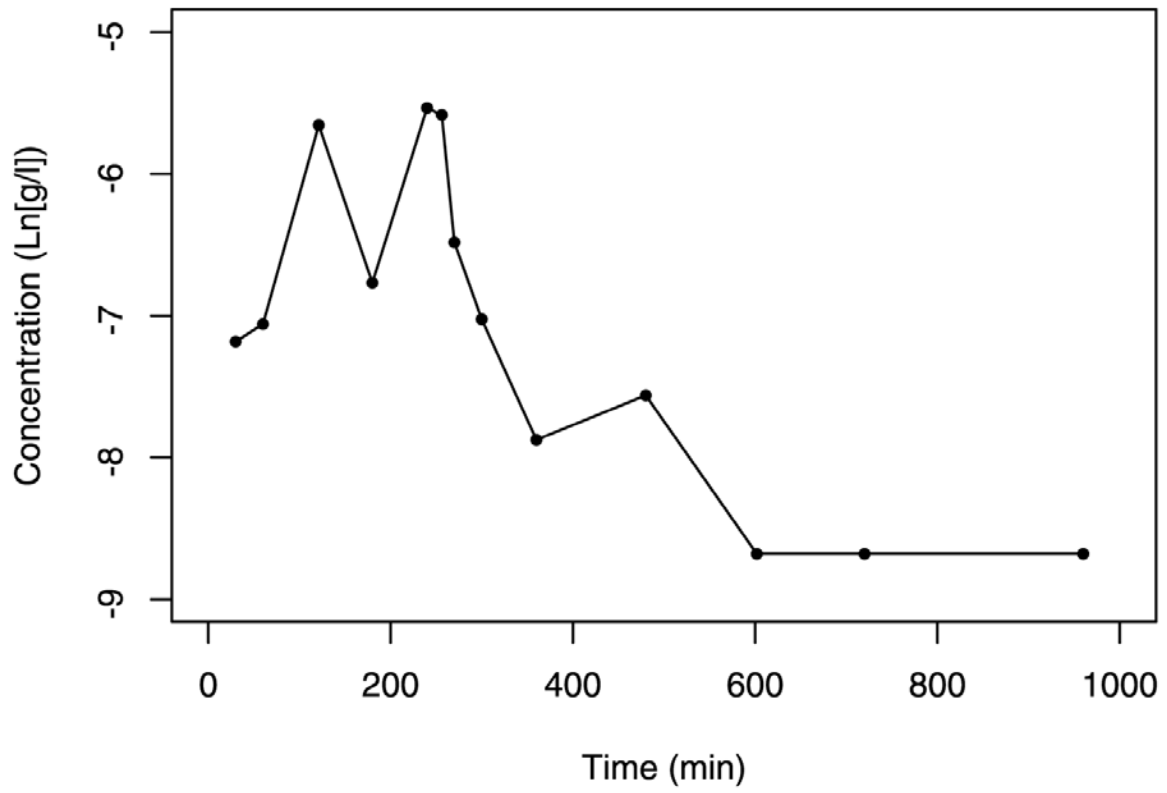


Figure 5: TCE concentration in venous blood generated by randomly sampling the data in Figure 1.

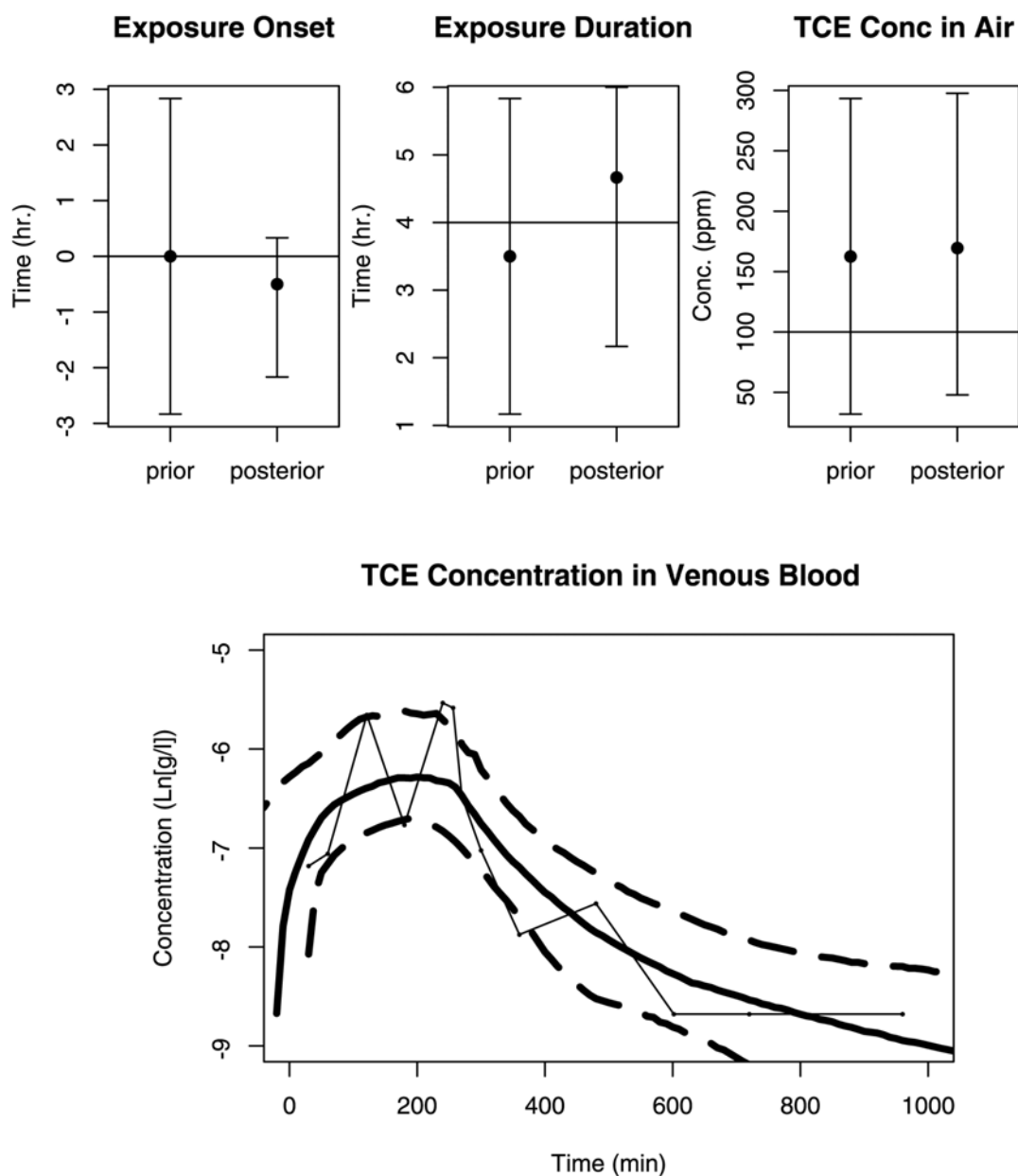


Figure 6: Reconstructed exposure profile and TCE concentration using the dataset in Figure 5. The prior and posterior bars in the barplot represents the uncertainty before and after comparison to biomarker data, respectively. The whiskers in the barplot represent the two-sided 95% confidence interval, the circle is the median, and the horizontal line represents the value reported by Fisher (1998). The thick dotted lines in the time series plot define the two-sided 95% confidence interval. The thick solid line is the median.